



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Faithful expression of a tagged Fugu WT1 protein from a genomic transgene in zebrafish

Citation for published version:

Miles, CG, Rankin, L, Smith, SI, Niksic, M, Elgar, G & Hastie, ND 2003, 'Faithful expression of a tagged Fugu WT1 protein from a genomic transgene in zebrafish: efficient splicing of pufferfish genes in zebrafish but not mice', *Nucleic Acids Research*, vol. 31, no. 11, pp. 2795-802.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Nucleic Acids Research

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Faithful expression of a tagged *Fugu* WT1 protein from a genomic transgene in zebrafish: efficient splicing of pufferfish genes in zebrafish but not mice

Colin G. Miles, Lesley Rankin, Shirley I. Smith, Martina Niksic, Greg Elgar¹ and Nicholas D. Hastie*

Medical Research Council Human Genetics Unit, Western General Hospital, Creve Road, Edinburgh EH4 2XU, UK and ¹United Kingdom Human Genome Mapping Project Resource Centre, Cambridge CB10 1SB, UK

Received February 19, 2003; Revised and Accepted April 3, 2003

ABSTRACT

The teleost fish are widely used as model organisms in vertebrate biology. The compact genome of the pufferfish, *Fugu rubripes*, has proven a valuable tool in comparative genome analyses, aiding the annotation of mammalian genomes and the identification of conserved regulatory elements, whilst the zebrafish is particularly suited to genetic and developmental studies. We demonstrate that a pufferfish *WT1* transgene can be expressed and spliced appropriately in transgenic zebrafish, contrasting with the situation in transgenic mice. By creating both transgenic mice and transgenic zebrafish with the same construct, we show that *Fugu* RNA is processed correctly in zebrafish but not in mice. Furthermore, we show for the first time that a *Fugu* genomic construct can produce protein in transgenic zebrafish: a full-length *Fugu* *WT1* transgene with a C-terminal β -galactosidase fusion is spliced and translated correctly in zebrafish, mimicking the expression of the endogenous *WT1* gene. These data demonstrate that the zebrafish:*Fugu* system is a powerful and convenient tool for dissecting both vertebrate gene regulation and gene function *in vivo*.

INTRODUCTION

The Wilms' tumour suppressor gene, *WT1*, was originally identified as a gene deleted or rearranged in Wilms' tumour (1,2), and *WT1* mutations have been shown to cause both Frasier (3) and Denys-Drash (4) syndromes. The gene encodes a four zinc finger-containing nuclear protein with roles both as a transcription factor and in RNA processing [reviewed in Englert (5)]. In addition to human genitourinary disorders, further evidence of the fundamental role of *WT1* in genitourinary development is provided by *WT1*-null mice, which display a complete absence of kidney and gonad structures and die in mid-gestation (6). A combination of alternative splicing, alternative translation initiation and RNA

editing gives rise to at least 16 distinct protein isoforms (5,7,8).

WT1 has been conserved throughout vertebrate evolution, although only mammals possess numerous protein isoforms (9,10). Non-mammalian vertebrates possess just two different protein isoforms resulting from an alternative splicing event which can insert three amino acids between the third and fourth zinc fingers. Given that the genitourinary system has evolved dramatically throughout vertebrate radiation, it is tempting to speculate that the complexity of mammalian *WT1* isoforms may be associated with changes to the genitourinary system, whilst the three amino acid alternative splice is involved in a basic function conserved through vertebrate evolution (11). The expression pattern of *WT1* has been conserved throughout vertebrates, with *WT1* expression being demonstrated in the developing genitourinary system of chick, alligator, *Xenopus* and zebrafish, as well as mammals (9,12–15). However, it should be noted that *WT1* is also expressed in genitourinary structures that are specific to mammals, such as the endometrium (16). Consequently, whilst evolutionary conservation may point to the basic function(s) of *WT1*, evolutionary differences could shed light on the acquisition of novel regulatory and functional properties during the evolution of an organ system.

The pufferfish, *Fugu rubripes*, has a genome that is approximately 7.5 times smaller than that of human, with compaction occurring within both introns and intergenic regions (17). The relatively small size of the *Fugu* genome has enabled a large amount of sequence data to be generated (18) (<http://fugu.hgmp.mrc.ac.uk>), thereby facilitating base-by-base comparison of fish and mammalian genomes (e.g. *Fugu*:human) and providing a framework onto which other genomes such as zebrafish and mouse may be superimposed. Typically, coding sequences are highly conserved between *Fugu* and mammals, and the randomisation of non-coding DNA sequences has led to the identification of regulatory elements associated with genes based upon non-coding DNA sequence conservation (19). Recent studies have demonstrated the usefulness of zebrafish transgenesis as a means of dissecting complex regulatory elements of *Fugu* genes (20). The strategy of using *Fugu* transgenes in zebrafish benefits from being a heterologous system, with the transgene readily

*To whom correspondence should be addressed. Tel: +44 131 332 2471; Fax: +44 131 467 8424; Email: nick.hastie@hgu.mrc.ac.uk

distinguishable from the endogenous gene, without the drawbacks of using evolutionarily distant species such as mouse in which high level tissue-specific expression of transgenes is not guaranteed (21).

Despite the widespread use of *Fugu* as a model organism for genome comparisons and the proven worth of utilising its compact genome to identify and characterise regulatory elements, entire *Fugu* genes have not been exploited for functional studies. The compact nature of *Fugu* genes suggests they would be a very useful source of genomic DNA constructs for transgenesis, with the advantage over cDNA constructs of being able to generate all isoforms of a particular gene, expressed from their own regulatory elements. The recently sequenced human genome (22,23) possesses fewer genes than predicted, and it has been suggested that alternative splicing may account for this discrepancy, resulting in a proteome of far greater complexity than the genome (24). Consequently, the ability to express all isoforms of a particular gene for functional studies is an important consideration.

To evaluate the potential of *Fugu* transgenes, Sathasivam *et al.* introduced the *Fugu* homologue of the Huntingdon's disease gene into a mouse cell line and created a line of transgenic mice harbouring a genomic *Fugu HD* construct (25). Despite detecting transgene expression, the *Fugu HD* transcript was aberrantly processed, with the inclusion of introns often giving rise to in-frame stop codons incompatible with production of full-length protein.

In this study, we have confirmed that *Fugu* RNAs are spliced incorrectly in transgenic mice, consistent with the aberrant splicing of *Fugu HD* observed by Sathasivam *et al.* Our data suggest there may be a general inability of *Fugu* RNAs to be processed correctly in mammalian cells due to inherent differences in splicing between fish and mammals. Consistent with this, we show that *Fugu* introns are spliced correctly in transgenic zebrafish and that the complete *Fugu WT1* gene is expressed in the correct tissue-specific manner under its own regulatory elements. Furthermore, this expression results in correct mRNA processing and translation leading to faithful expression of the heterologous protein in transgenic zebrafish, demonstrating for the first time that *Fugu* genomic transgenes can be used to express *Fugu* proteins in transgenic zebrafish.

MATERIALS AND METHODS

DNA constructs

WT1-containing cosmids were isolated from a gridded cosmid library of *F. rubripes* genomic DNA [Greg Elgar, Human Genome Mapping Project Resource Centre (HGMP-RC)], as described (10). Unmodified constructs for mouse transgenesis were isolated directly from *Fugu* cosmids by restriction digestion. *Fugu RCN1* (Fig. 1A) was isolated as an 18 kb *HindIII* fragment from cosmid 151J19, and *Fugu WT1* (Fig. 1C) as a 12 kb *BamHI*-*XbaI* fragment from cosmid 104M11.

The splicing construct HAFrWT1MYC was created by cloning exons 5–9 of *Fugu WT1* from 151J19 (Fig. 1B) into a modified pCI-neo (Promega) carrying three N-terminal haemagglutinin (HA) tags and a C-terminal myc tag (26).

For microinjection into mouse oocytes, the transgene was liberated from vector sequences by *BglIII*-*NaeI* digestion.

pXIG (a kind gift from Adam Amsterdam) consists of the green fluorescent protein (GFP)-coding sequence under control of the *Xenopus* elongation factor 1 α promoter. This construct directs ubiquitous GFP expression in transgenic zebrafish (27) and was injected in its supercoiled form to obtain transient expression in zebrafish embryos.

To create pFrWTlac for fish transgenesis, *Fugu WT1* was first cloned into pPolyIII as a 12 kb *BamHI*-*XbaI* fragment (Fig. 1C). A β -galactosidase gene, stop and polyadenylation sequences were cloned into the *SphI* site in exon 9, using oligonucleotide linkers to create a fusion construct, retaining the open reading frame and allowing release of the construct by *NotI* digestion.

Production of transgenic animals

Constructs for transgenic mice were digested using the restriction enzymes described above, gel purified to remove all vector sequences and injected into fertilised (CBA \times C57Bl/6)F₂ oocytes at ~5 ng/ μ l. Positive founder mice were identified by Southern blot, bred with (CBA \times C57Bl/6)F₁ mice and lines were maintained as heterozygotes. Transient transgenic embryos were harvested 10 days after oviduct transfer of microinjected oocytes into pseudopregnant recipients and genotyped by PCR of embryonic tail DNA with transgene-specific primers.

Transgenic zebrafish were obtained by injecting single blastomeres of one- and two-cell stage zebrafish embryos (Gol^{+/−}) with pHAFrWT1MYC at ~15 ng/ μ l. Co-injection with the ubiquitously expressing GFP vector, pXIG (27), enabled selection of transgenic (expressing) embryos.

Analysis of transgenic animals

Total RNA was isolated from mouse tissues, whole mouse embryos and zebrafish embryos using the RNagents isolation system (Promega), then subjected to DNase I treatment (Roche Diagnostics) and a second phenol/chloroform extraction and precipitation. *Fugu* RNA was obtained from the MRC HGMP-RC and subjected to DNase I treatment (Roche Diagnostics) prior to reverse transcription. For DNA amplification, *Fugu* RNA was subjected to PCR without reverse transcription and prior to DNase I treatment. RT-PCR was performed either by using the Access RT-PCR kit (Promega) or by reverse transcribing total RNA using an oligo(dT) primer followed by 30 cycles of intron-specific PCR. To control for differences in amplification of particular PCR products during thermal cycling and preferential amplification due to variations in template concentration, PCRs were also carried out over a range of cycle numbers (24, 26, 28 and 30 cycles) or with varying amounts of starting template. Control RT-PCRs were performed to normalise loading and verify DNase I treatment using primers for *HPRT* or *RCN1* (mouse) and *cyclin D* (zebrafish). The identity of correctly sized RT-PCR products was confirmed by Southern blot and hybridisation with [γ -³²P]ATP-labelled internal oligonucleotides (oligonucleotide sequences available on request).

For X-gal staining, zebrafish embryos were fixed for 45 min in 4% paraformaldehyde (PFA) at room temperature, dechorionated in phosphate-buffered saline (PBS), washed in PBS containing 0.1% Tween 20 (PBT) and stained in X-gal

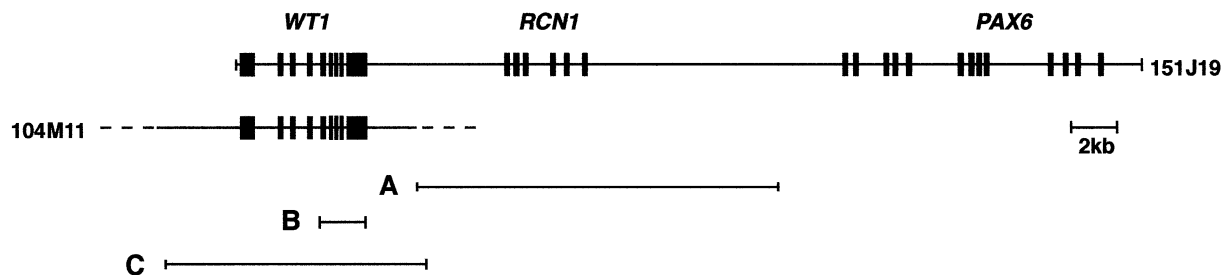


Figure 1. Scale drawing of the *Fugu rubripes* WAGR region encompassing *WT1*, *RCN1* and *PAX6* genes depicting the constructs used in this study. (A) *Fugu RCN1*. (B) PHAFrWT1MYC. (C) *Fugu WT1/pFrWTlac*.

stain solution [$\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, NaCl, $\text{K}_4\text{Fe}_3(\text{CN})_6/\text{K}_3\text{Fe}(\text{CN})_6$ plus 0.7 mg/ml X-gal] at 37°C for between 1 and 6 h.

Zebrafish whole-mount *in situ* hybridisation was performed essentially as described (28) using the full-length zebrafish *WT1* cDNA (GenBank accession no. AF144550) as a probe. Hybridisation of the digoxigenin-labelled riboprobe was detected using anti-DIG Fab-AP (1:5000) in combination with NBT/BCIP substrates (Roche). Embryos were equilibrated to 80% glycerol for photography.

RESULTS

Aberrant splicing of *Fugu* genes in transgenic mice

The compact nature of the *Fugu* genome suggests that *Fugu* transgenes may provide natural minigenes for transgenesis to study both gene regulation and gene function. We have previously cloned the *Fugu* equivalent of human chromosome 11p13, the WAGR region, and demonstrated that this locus is highly conserved between fish and mammals; gene order and orientation are conserved, as are numerous non-coding regions, such as known *Pax6* regulatory elements (10). This conservation implies that the *Fugu* WAGR region is a good model for its mammalian equivalent, and the existence of mouse strains carrying a range of deletions across this region (29,30) makes it an attractive locus at which to perform transgenic rescue experiments in mice.

As a first step towards this goal, we generated eight lines of transgenic mice via pronuclear injection of either *Fugu RCN1* or *Fugu WT1* into fertilised, wild-type, F_1 oocytes. Two out of the five lines of *Fugu RCN1* transgenic mice displayed widespread transgene expression, consistent with the expression pattern of endogenous *RCN1*. Transgene expression was only detectable in one line of *Fugu WT1* transgenic mice, with expression restricted to the thymus (not shown). As endogenous *WT1* is not expressed in the thymus, this represents a site of ectopic transgene expression, suggesting either that the *Fugu WT1* transgene lacks important regulatory elements or that the species differences between pufferfish and mammals are too great for reliable *Fugu WT1* expression in transgenic mice.

We performed RT-PCR to determine whether the compact *Fugu* transgenes were spliced correctly and thus were capable of generating full-length protein. Consistent with previous transgenic mice expressing *Fugu HD* (25), aberrant splicing was observed in all three expressing lines of transgenic mice.

To investigate this aberrant splicing in detail, RT-PCR was performed on a panel of tissues from the highest expressing *Fugu RCN1* line across specific intron–exon boundaries (Fig. 2). *Fugu RCN1* introns of <200 bp were not spliced in transgenic mice (introns 2 and 5), whilst the 1.22 kb intron 1 was fully spliced. Introns 3 and 4 (1.23 kb and 250 bp, respectively), however, were partly spliced. In contrast, RT-PCR analysis of *Fugu* RNA showed that the endogenous *RCN1* RNA was fully spliced *in vivo*, with no evidence of unspliced RNA even at limiting cycle number. The doublet arising from amplification of intron 4 is likely to be due to a pseudogene which has been identified elsewhere in the *Fugu* genome. It has been suggested that *Fugu* introns are too small to be spliced efficiently by the mammalian splicing machinery (25). The observation that *Fugu RCN1* intron 1 (1.22 kb) is fully spliced whilst intron 3 (1.23 kb) is poorly spliced implies that intron size alone is not sufficient to account for this phenomenon. Consistent with this, mammalian cells are capable of efficiently splicing introns of <100 bp, e.g. intron 17 of human *HD* is 93 bp (31).

Efficient splicing of *Fugu* introns in transgenic zebrafish

Given that species differences rather than intron size were the likely cause of aberrant splicing in *Fugu* transgenic mice, we tested whether *Fugu* introns could be processed correctly in a transgenic fish system.

To facilitate direct comparison between transgenic mice and transgenic zebrafish, a cytomegalovirus (CMV)-driven *Fugu WT1* splicing construct was utilised. This construct, pHAfrWT1MYC, consisted of a CMV promoter driving expression of a genomic fragment encompassing *Fugu WT1* exons 5–9 (Fig. 1B), with in-frame N- and C-terminal tags to facilitate specific detection of transgenic transcripts by RT-PCR. The *Fugu WT1* introns within this construct range in size from 80 to 500 bp (Fig. 3A).

We generated 15 transient transgenic mouse embryos with pHAfrWT1MYC. Microinjected embryos were taken at e10.5, genotyped by PCR of DNA extracted from the embryo tail and total RNA was isolated from individual whole embryos. In all transient transgenic mouse embryos, inefficient splicing was observed across individual introns and, when splicing was analysed across two introns, almost no correctly spliced transcripts were detectable. Figure 3B shows RT-PCR analysis of splicing across exons 7–9 for two transgenic mice, M10 and M14, representative of low and high expressors, respectively. Given the inefficiency of splicing across two introns, it is unlikely that any correctly spliced

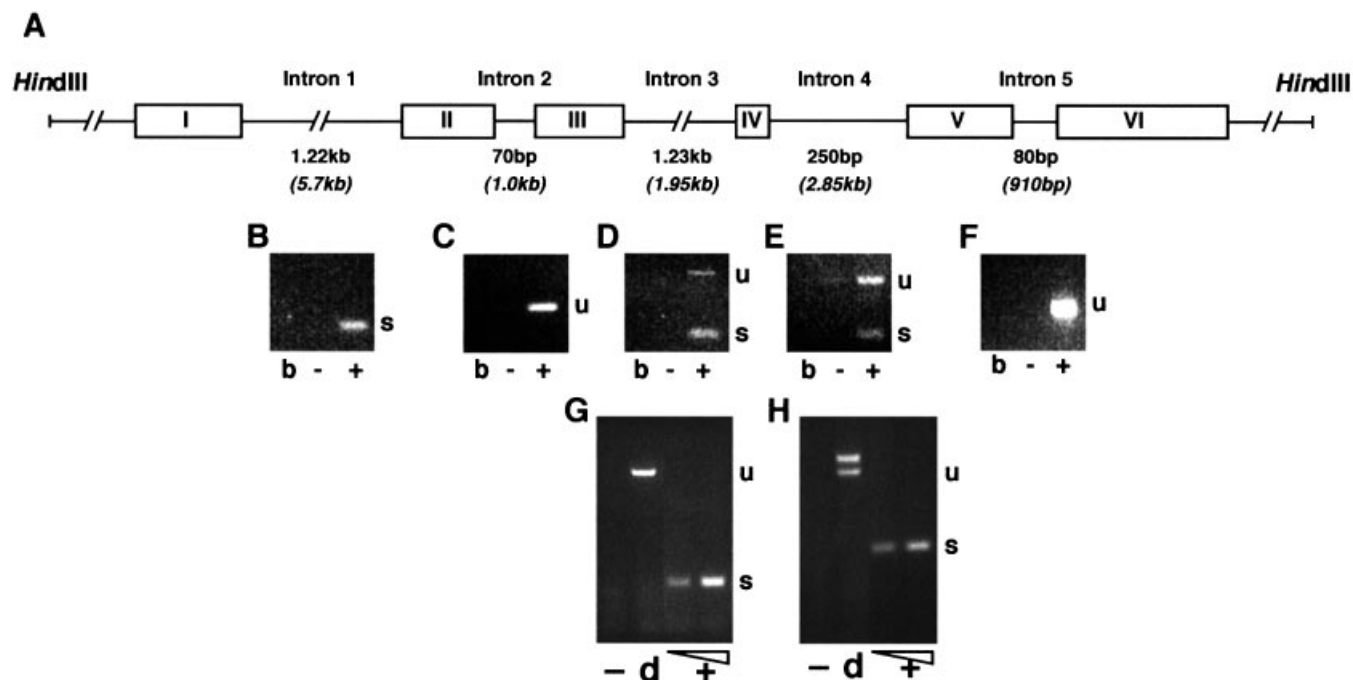


Figure 2. A *Fugu RCN1* transgene is not spliced efficiently in transgenic mice. (A) Diagram of the *Fugu RCN1* genomic transgene showing intron sizes. The sizes of the corresponding mouse introns are shown in parentheses. (B–F) RT–PCR analysis of splicing for individual introns of *Fugu RCN1* in transgenic mice. (B) Intron 1, (C) intron 2, (D) intron 3, (E) intron 4, (F) intron 5. Total RNA prepared from kidney of transgenic *Fugu RCN1* mice was subjected to RT–PCR across individual intron–exon boundaries. Lane b, H₂O blank. (G and H) RT–PCR analysis of the endogenous *Fugu RCN1* RNA in pufferfish kidney. Lane d, DNA amplification; negative control (–) consisting of an equivalent amount of the total RNA preparation subjected to PCR without reverse transcription; PCR of reverse-transcribed RNA (+). RT–PCR products corresponding to spliced and unspliced introns are labelled s and u, respectively.

full-length mRNA would be produced in transgenic mice harbouring a complete *Fugu WT1* gene. The splicing defects observed in HAFrWT1MYC transgenic embryos were not restricted to intron inclusion, with exon skipping being observed in highly expressing individuals.

The skipping of exon 8 and the correct splicing of exons 7–9 of *Fugu WT1* were only detectable by conventional RT–PCR at low levels in highly expressing transgenic mice, suggesting that these are rare events and that intron inclusion is the most common splicing defect when *Fugu* genes are expressed in mouse cells. The inclusion of either intron 7 or intron 8 would result in four (intron 7) or three (intron 8) in-frame translational stop codons within the aberrantly spliced mRNA.

To directly compare splicing efficiency of this transgene between mice and fish, we generated transgenic zebrafish via co-injection of pHAFrWT1MYC and pXIG. Transgenic embryos were selected for green fluorescence at 10–12 hours post-fertilisation (h.p.f.), pooled in groups of between 15 and 30, and total RNA was isolated from separate pools for RT–PCR analysis. Two representative pools of transgenic zebrafish embryos are shown in Figure 3B. In contrast to transgenic mouse embryos, splicing of the transgene in zebrafish embryos was far more efficient. To control for saturation effects in the PCRs, ZF28 and M10 samples were subject to PCR for between 24 and 30 cycles (Fig. 3C), showing that the aberrantly spliced RNA can be detected in transgenic mouse embryos at limiting cycle number and that no unspliced RNA is detected in transgenic zebrafish embryos. Thus the same *Fugu* construct is spliced efficiently in

zebrafish but not in mice, suggesting that the differences affecting splicing are not specific to the pufferfish but reflect species differences between fish and mammals in general.

The alternative splicing of three amino acids between zinc fingers 3 and 4 of *WT1* is crucial for *WT1* function within the genitourinary system and is conserved throughout all vertebrates (9–11). This alternative splicing event is not regulated by a tissue-specific factor and even occurs in transfected cells not expressing endogenous *WT1* (26). Given the efficiency with which the *Fugu WT1* construct was spliced in transgenic zebrafish, we used RT–PCR to determine whether the transgene was also alternatively spliced. RT–PCR analyses from exon 8 to exon 9 on pools of transgenic zebrafish demonstrate that the *Fugu WT1* construct is alternatively spliced (Fig. 3D) and that *Fugu* genomic transgenes have the potential to give rise to the full range of alternatively spliced *WT1* transcripts in zebrafish. We also analysed pHAFrWT1MYC transgenic mouse embryos in which low levels of correctly spliced product were observed using RT–PCR from exon 8 to exon 9 but with the addition of [³²P]dCTP to the reaction. By increasing sensitivity in this way, we found that the *Fugu* construct also exhibits alternative splicing in transgenic mice (data not shown), consistent with the ubiquitous nature of this alternative splicing event.

Protein expression from a full-length *Fugu WT1* transgene in zebrafish

The efficiency with which the *Fugu* genomic construct is spliced in fish suggests that full-length *Fugu* mRNA can be

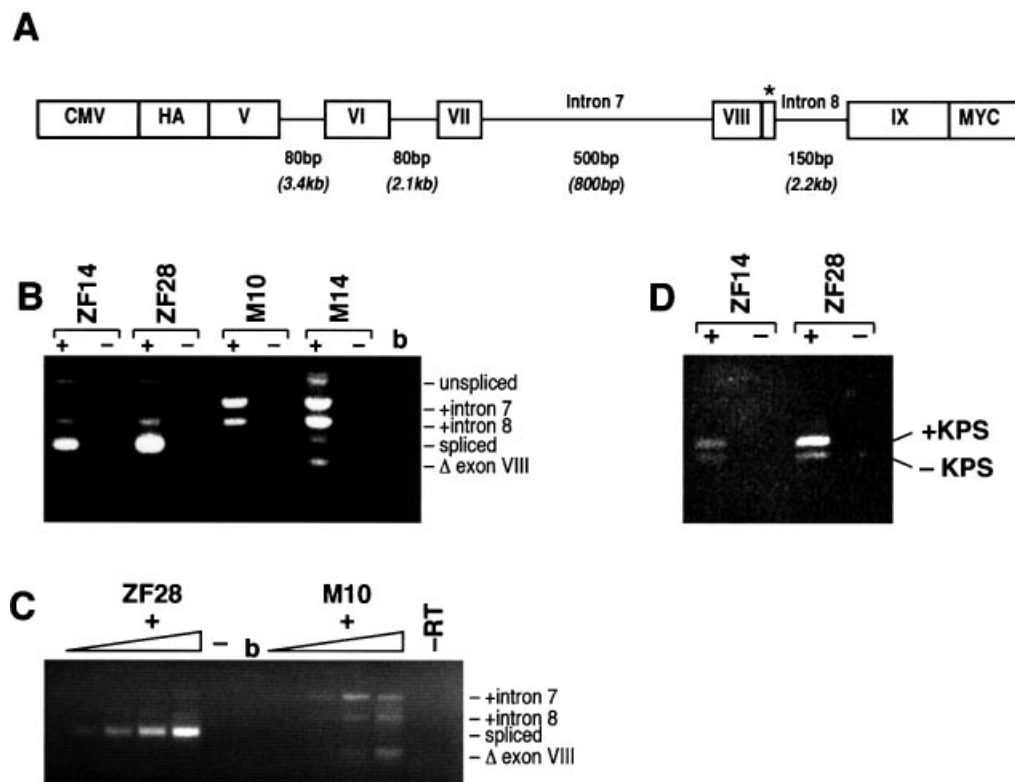


Figure 3. Efficient splicing of a *Fugu rubripes* transgene in zebrafish but not in mice. (A) Diagram of pHAFrWT1MYC showing intron sizes, with the sizes of the corresponding mouse introns in parentheses. The asterisk refers to the location of the three amino acid alternative splicing event. Note that fish do not possess an equivalent to mammalian *WT1* exon 5. Consequently, *Fugu WT1* exon 6 corresponds to mammalian exon 7 and so on. (B) RT-PCR analysis of transgenic zebrafish pools, ZF14 and ZF28, and individual transgenic mouse embryos, M10 and M14, selected to represent low and high expressors respectively. RT-PCR was performed on total RNA isolated from whole embryos using oligonucleotides situated in exon 7 and the myc tag. The negative control consists of an equivalent amount of the total RNA preparation subjected to PCR without reverse transcription, b = H₂O blank. RT-PCR products corresponding to each partly spliced transcript are labelled. (C) RT-PCR analysis of ZF28 and M10 RNA from exon 7 to MYC showing the emergence of each class of PCR product from 24 to 30 cycles. An equivalent amount of RNA was amplified for 30 cycles without reverse transcription as a control (–RT), b = H₂O blank. (D) Analysis of alternative splicing of the transgene in pools of transgenic zebrafish (ZF14 and ZF28) harbouring pHAFrWT1MYC. RT-PCR was performed with oligonucleotides situated in exon 8 and the myc tag. The negative control (–) consists of an equivalent amount of the total RNA preparation subjected to PCR without reverse transcription.

produced in transgenic zebrafish. To test this directly and to determine whether *Fugu* mRNA can be translated in zebrafish to produce full-length protein, we modified the complete *Fugu WT1* gene and flanking genomic regions to include a histochemical marker (β -galactosidase) fused to the penultimate codon of *WT1*, such that only faithful splicing and translation of the resulting fusion mRNA would produce a fusion protein detectable by X-gal staining (Fig. 4A).

We injected pFrWTlac into fertilised zebrafish eggs. Approximately 10% of injected embryos displayed X-gal staining, demonstrating that the full-length genomic *Fugu WT1* transgene was indeed correctly spliced and translated to produce full-length protein. Injected embryos were stained between 10 and 72 h.p.f. At 24 h.p.f., the degree of mosaicism associated with zebrafish transgenesis resulted in expression being restricted to small ‘clones’ of cells (data not shown). At 12 h.p.f., it was possible to identify transgenic embryos with more complete expression patterns—typically within the developing pronephros (Fig. 4C)—although individual embryos exhibited a great deal of mosaicism (compare Fig. 4D and E). The expression pattern of the *Fugu* transgene observed at 12 h.p.f. recapitulated that of endogenous

zebrafish *WT1* (Fig. 4B). X-gal staining of zebrafish embryos injected with pFrWTlac at later time points (48 and 72 h.p.f.) revealed no transgenic fish, with numerous embryos dying or developing abnormally. By performing zebrafish injections with a modified pFrWTlac (pFrWTGFP) containing GFP in place of lacZ, we confirmed that expression of the transgene, as determined by GFP expression in unfixed zebrafish embryos, was associated with abnormal development and lethality (not shown). This finding is consistent with the suggestion that overexpression of *WT1* in transgenic mice causes embryonic lethality *in vivo* (32).

DISCUSSION

In this study, we demonstrate that *Fugu* transgenes are aberrantly spliced in transgenic mice and yet are correctly spliced in transgenic zebrafish. Furthermore, we show that *Fugu WT1* is expressed faithfully in the developing kidney of transgenic zebrafish, resulting in the production of a marked *Fugu WT1* protein from the transgene. These observations broaden the utility of *Fugu* as a model system to include

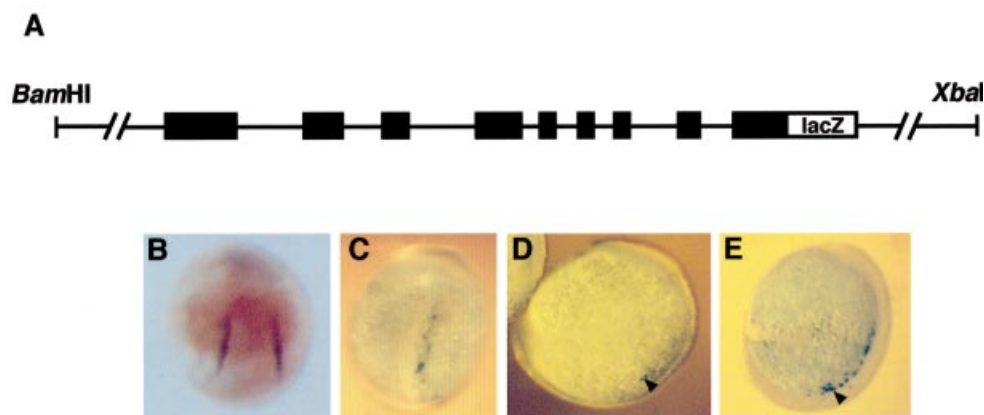


Figure 4. Expression of full-length protein from a pufferfish transgene in zebrafish. (A) Diagram of the genomic *Fugu WT1* fragment with C-terminal β -galactosidase fusion. (B) Expression of endogenous *WT1* in control 12 h.p.f. zebrafish by *in situ* hybridisation. (C) X-gal-stained transgenic zebrafish at ~12 h.p.f. Correct splicing of the construct is evident from X-gal staining resulting from production of a fusion protein. The mosaicism associated with zebrafish transgenesis results in staining only in one half of the pronephric field. (D and E) Examples of mosaicism observed in pronephric expression of *Fugu WT1-lacZ* fusion protein between individual fish. Arrowheads point to transgene expression (X-gal staining) in the developing pronephros.

functional experiments *in vivo* using *Fugu* genes in transgenic zebrafish.

***Fugu* introns are spliced efficiently in transgenic zebrafish but not transgenic mice**

The failure of *Fugu* transgenes to be spliced accurately in mammalian cells is consistent with a previous study of *Fugu HD* in which Sathasivam *et al.* speculate that the small size of *Fugu* introns stresses the fidelity of the mammalian splicing apparatus (25). Whilst our data support this, they also suggest that the size of a particular intron is not the only factor determining whether that intron is spliced efficiently. Consistent with this, many mammalian introns are <250 bp (33).

A survey of more than 100 splice site mutations found that the most common consequence of splice site mutation was exon skipping, with intron inclusion only observed in 6% of cases (34). In contrast, intron inclusion is the most common splicing defect observed when *Fugu* genes are expressed in mouse cells. Therefore, it seems unlikely that specific sequence differences can fully explain the aberrant splicing of *Fugu* constructs in mouse cells, as such differences might be expected to result in exon skipping. In addition, the fact that individual *Fugu* introns can be partially spliced suggests that this phenomenon is related to evolutionary differences at the level of splicing efficiency between fish and mammals, in general. This implies that *Fugu* resembles other teleosts at the level of splicing despite its compact genome. It may be that these evolutionary differences between splicing in fish and mammals could be exploited to further understand the fundamental mechanisms underlying splicing in vertebrates.

Typically, vertebrate genes possess large introns and small exons, whilst the genes of lower eukaryotes possess small introns and relatively large exons (35). These differences form the basis of the intron versus exon definition models of splicing (36). *Fugu*, however, is an exception in that its genes display vertebrate exon organisation with generally small introns (17). This reduction in intron size whilst retaining exon size compared with other vertebrates may, as these models predict, stress the fidelity of the splicing machinery of mouse

cells. Intriguingly, the observation that *Fugu* introns are spliced efficiently in zebrafish would suggest general species differences in intron versus exon definition between fish and mammals, rather than a phenomenon unique to *Fugu*. Cross-species transgenesis may provide an interesting way in which to study these fundamental splicing decisions which, to date, have been addressed by making changes to a splicing construct within a constant cell system. By keeping the splicing construct constant and changing the cell (species) type, parameters such as ionic strength, factor concentration and temperature may be affecting the activity of splicing factors or the configuration of pre-mRNA (37–39). In light of this, it is interesting to note the recent characterisation of an alternative splicing event within the human neurofibromatosis type 1 mRNA which is induced by cold stress of 25°C (40).

The genomes of both the pufferfish (<http://fugu.hgmp.mrc.ac.uk>) and the zebrafish (http://www.sanger.ac.uk/Projects/D_rerio) are in the process of being fully sequenced. The complete genome sequence of these two teleost fish will enable detailed genomic comparisons to be made between fish and mammals, and will complement cross-species transgenic experiments of the kind described herein. In particular, the comparison of transgenic mice harbouring zebrafish genomic clones with mice harbouring pufferfish genomic constructs would address the issue of whether these splicing differences truly are teleost specific and also serve as a valuable experimental resource for investigating the mechanism of such differences.

Faithful expression of a marked *Fugu WT1* protein in transgenic zebrafish

The sequences regulating the expression of *WT1* *in vivo* have yet to be identified. The proximal promoter of the human gene directs variable expression of a reporter gene in transgenic mice and not in the major site of *WT1* expression, the developing genitourinary system (41). A 280 kb human *WT1* yeast artificial chromosome (YAC) is required to achieve correct genitourinary expression of a reporter gene in transgenic mice (41,42) although expression of this construct is influenced by its site of integration. Despite the

benefits of recombination in yeast to modify constructs, YAC transgenesis as a means of dissecting gene regulation is costly and time consuming. The conservation of genitourinary *WT1* expression throughout vertebrates suggests that gene regulation studies in lower vertebrates and comparative evolutionary approaches may be appropriate for *WT1*. The pufferfish genome is approximately 7.5 times smaller than that of human (17), whilst the *Fugu WT1* gene is approximately 15 times smaller than its human counterpart (10). We demonstrate that, despite this compaction, the *Fugu WT1* gene can faithfully recapitulate the expression pattern of the endogenous zebrafish *WT1* gene when introduced as a transgene.

The inability to detect expression of *Fugu WT1* in the kidneys of transgenic mice may be due to the mammalian transcription machinery not recognising *Fugu WT1* elements efficiently or, as with the analogous human construct, may reflect a great sensitivity to site of integration effects. That a similar *Fugu* construct directs reproducible kidney-specific expression in transgenic zebrafish further strengthens the case for the *Fugu:zebrafish* system as a means of dissecting gene regulation *in vivo*, and provides a genetic tool for studying pronephric development in zebrafish via transgenesis.

Whilst this study confirms that *Fugu* genes do not function as natural minigenes in transgenic mice, full-length *Fugu WT1* with a C-terminal β -galactosidase fusion is correctly spliced and translated to produce the fusion protein in transgenic zebrafish. These studies suggest that *Fugu* genes will be useful for functional transgenic experiments in zebrafish. For example, *Fugu* cosmid introduced into zebrafish mutants could be employed in transgenic rescue experiments—*Fugu* genes can express faithfully in zebrafish, *Fugu* mRNAs can be processed to give protein and *Fugu* clones are available ‘off the shelf’ from the UK HGMP-RC (<http://fugu.hgmp.mrc.ac.uk>). Given the increasing number of zebrafish mutants available, this approach could have broad implications. The zebrafish:*Fugu* system now represents a powerful and convenient tool for dissecting both vertebrate gene regulation and gene function *in vivo* and may help in the study of the fundamental mechanisms underlying vertebrate splicing.

ACKNOWLEDGEMENTS

Thanks to Adam Amsterdam and Rachel Davies for plasmids pXIG and pHAfrWT1MYC, respectively. We are grateful to the Currie laboratory for discussions and the use of the zebrafish facility and thank Rachel Brough for excellent technical assistance. Thanks also to the HGU Photographics Department. This work was funded by the Medical Research Council (UK). C.G.M. was in receipt of an MRC Research (Training) Fellowship, and S.I.S. is a Howard Florey Fellow.

REFERENCES

- Call, K.M., Glaser, T., Ito, C.Y., Buckler, J., Pelletier, J., Haber, D.A., Rose, E.A., Kral, A., Yeger, H. and Lewis, W.H. (1990) Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell*, **60**, 509–520.
- Gessler, M., Poustka, A., Cavenee, W., Neve, R.L., Orkin, S.H. and Bruns, G.A. (1990) Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. *Nature*, **343**, 774–778.
- Barboux, S., Niaudet, P., Gubler, M.C., Grunfeld, J.P., Jaubert, F., Kuttann, F., Fekete, C.N., Souleyreau, T.N., Thibaud, E., Fellous, M. and McElreavey, K. (1997) Donor splice mutations in *WT1* are responsible for Frasier syndrome. *Nature Genet.*, **17**, 467–470.
- Pelletier, J., Bruening, W., Kashtan, C.E., Mauers, S.M., Manivel, J.C., Striegel, J.E., Houghton, D.C., Junien, C., Habib, R., Fouser, L. *et al.* (1991) Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys–Drash syndrome. *Cell*, **67**, 437–447.
- Englert, C. (1998) *WT1*—more than a transcription factor? *Trends Biochem. Sci.*, **23**, 389–393.
- Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D. and Jaenisch, R. (1993) *WT1* is required for early kidney development. *Cell*, **74**, 679–691.
- Hastie, N.D. (1994) Wilms' tumour: a case of disrupted development. *Annu. Rev. Genet.*, **28**, 523–558.
- Lee, S.B. and Haber, D.A. (2001) Wilms tumor and the *WT1* gene. *Exp. Cell Res.*, **264**, 74–99.
- Kent, J., Coriat, A.M., Sharpe, P.T., Hastie, N.D. and Van Heyningen, V. (1995) The evolution of *WT1* sequence and expression pattern in the vertebrates. *Oncogene*, **11**, 1781–1792.
- Miles, C., Elgar, G., Coles, E., Kleinjan, D.J., Van Heyningen, V. and Hastie, N. (1998) Complete sequencing of the *Fugu* WAGR region from *WT1* to *PAX6*: dramatic compaction and conservation of synteny. *Proc. Natl Acad. Sci. USA*, **95**, 13068–13072.
- Hastie, N.D. (2001) Life, sex and *WT1* isoforms—three amino acids can make all the difference. *Cell*, **106**, 391–394.
- Semba, K., Saito-Ueno, R., Takayama, G. and Kondo, M. (1996) cDNA cloning and its pronephros-specific expression of the Wilms' tumour suppressor gene, *WT1*, from *Xenopus laevis*. *Gene*, **175**, 167–172.
- Drummond, I.A., Majumdar, A., Hentschel, H., Elger, M., Solnica-Krezel, L., Schier, A.F., Neuhauss, S.C., Stemple, D.L., Zwartkruis, F., Rangini, Z. *et al.* (1998) Early development of the zebrafish pronephros and analysis of mutations affecting pronephric function. *Development*, **125**, 4655–4667.
- Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., Housman, D. *et al.* (1990) The candidate Wilms' tumour gene is involved in genitourinary development. *Nature*, **346**, 194–198.
- Armstrong, J.F., Pritchard-Jones, K., Bickmore, W.A., Hastie, N.D. and Bard, J.B. (1993) The expression of the Wilms' tumour gene, *WT1*, in the developing mammalian embryo. *Mech. Dev.*, **40**, 85–97.
- Zhou, J., Rauscher, F.J., III and Bondy, C. (1993) Wilms' tumor (*WT1*) gene expression in rat decidua differentiation. *Differentiation*, **54**, 109–114.
- Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. and Aparicio, S. (1993) Characterisation of the pufferfish (*Fugu*) genome as a compact model genome. *Nature*, **366**, 265–268.
- Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.K., Bouchireb, N., Cottage, A., Yeo, G.S.H., Umrana, Y. *et al.* (1999) Generation and analysis of 25Mb of genomic DNA from the pufferfish (*Fugu rubripes*) by sequence scanning. *Genome Res.*, **9**, 960–971.
- Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P., Krumlauf, R. and Brenner, S. (1995) Detecting conserved regulatory elements with the model genome of the Japanese pufferfish, *Fugu rubripes*. *Proc. Natl Acad. Sci. USA*, **92**, 1684–1688.
- Barton, L.M., Götting, B., Gering, M., Gilbert, J.G.R., Grafham, D., Rogers, J., Bentley, D., Patient, R. and Green, A.R. (2001) Regulation of the stem cell leukemia (*SCL*) gene: a tale of two fishes. *Proc. Natl Acad. Sci. USA*, **98**, 6747–6752.
- Nonchev, S., Maconochie, M., Vesques, C., Aparicio, S., Ariza-McNaughton, L., Manzanarez, M., Maruthinar, K., Kuroiwa, A., Brenner, S., Charnay, P. *et al.* (1996) The conserved role of *Krox-20* in directing Hox gene expression during hindbrain segmentation. *Proc. Natl Acad. Sci. USA*, **93**, 9339–9345.
- International Human Genome Sequencing Consortium (2001) The sequence of the human genome. *Nature*, **409**, 860–921.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304–1351.
- Modrek, B. and Lee, C. (2002) A genomic view of alternative splicing. *Nature Genet.*, **30**, 13–19.
- Sathasivam, K., Baxendale, S., Mangiarini, L., Bertaux, F., Hetherington, C., Kanazawa, I., Lehrach, H. and Bates, G.P. (1997) Aberrant processing of the *Fugu* HD (FrHD) mRNA in mouse cells and in transgenic mice. *Hum. Mol. Genet.*, **6**, 2141–2149.

26. Davies,R.C., Bratt,E. and Hastie,N.D. (2000) Did nucleotides or amino acids drive evolutionary conservation of the WT1+/- KTS alternative splice? *Hum. Mol. Genet.*, **9**, 1177–1183.
27. Amsterdam,A., Lin,S. and Hopkins,N. (1995) The *Aequorea victoria* green fluorescent protein can be used as a reporter in live zebrafish embryos. *Dev. Biol.*, **171**, 123–129.
28. Broadbent,J. and Read, E.M. (1999) Wholemount *in situ* hybridization of *Xenopus* and zebrafish embryos. *Methods Mol. Biol.*, **127**, 57–67.
29. Glaser,T., Lane,J. and Housman,D. (1990) A mouse model of the aniridia-Wilms' tumor deletion syndrome. *Science*, **250**, 823–827.
30. Kent,J., Lee,M., Schedl,A., Boyle,S., Fantes,J., Powell,M., Rushmere,N., Abbott,C., Van Heyningen,V. and Bickmore,W. (1997) The reticulocalbin gene maps to the WAGR region in human and the Small eye Harwell deletion in mouse. *Genomics*, **42**, 260–267.
31. Baxendale,S., Abdulla,S., Elgar,G., Buck,D., Berks,M., Mickelm,G., Durbin,R., Bates,G., Brenner,S., Beck,S. *et al.* (1995) Comparative sequence analysis of the human and pufferfish *Huntingdon's* disease genes. *Nature Genet.*, **10**, 67–75.
32. Menke,A., McInnes,L., Hastie,N.D. and Schedl,A. (1998) The Wilms' tumor suppressor WT1: approaches to gene function. *Kidney Int.*, **53**, 1512–1518.
33. Elgar,G., Sandford,R., Aparicio,S., Macrae,A., Venkatesh,B. and Brenner,S. (1996) Small is beautiful: comparative genomics with the pufferfish (*Fugu rubripes*). *Trends Genet.*, **12**, 145–150.
34. Nakai,K. and Sakamoto,H. (1994) Construction of a novel database containing aberrant splicing mutations of human genes. *Gene*, **141**, 171–177.
35. Sterner,D.A., Carlo,T. and Berget,S.M. (1996) Architectural limits on split genes. *Proc. Natl Acad. Sci. USA*, **93**, 15081–15085.
36. Berget,S.M. (1995) Exon recognition in vertebrate splicing. *J. Biol. Chem.*, **270**, 2411–2414.
37. Reed,R. and Maniatis,T. (1986) A role for exon sequences and splice-site proximity in splice-site selection. *Cell*, **46**, 681–690.
38. Schmitt,P., Gattoni,R., Keohavong,P. and Stevenin,J. (1987) Alternative splicing of E1A transcripts of adenovirus requires appropriate ionic conditions *in vitro*. *Cell*, **50**, 31–39.
39. Weil,D., D'Alessio,M., Ramirez,F., Steinmann,B., Wirtz,M.K., Glanville,R.W. and Hollister,D.W. (1989) Temperature-dependent expression of a collagen splicing defect in the fibroblasts of a patients with Ehlers–Danlos syndrome type VII. *J. Biol. Chem.*, **264**, 16804–16809.
40. Ars,E., Serra,E., de la Luna,S., Estivill,X. and Lázaro,C. (2000) Cold shock induces the insertion of a cryptic exon in the neurofibromatosis type 1 (NF1) mRNA. *Nucleic Acids Res.*, **28**, 1307–1312.
41. Moore,A.W., McInnes,L., Kreidberg,J., Hastie,N.D. and Schedl,A. (1999) YAC complementation shows a requirement for Wt1 in the development of epicardium, adrenal gland and throughout nephrogenesis. *Development*, **126**, 1845–1857.
42. Moore,A.W., Schedl,A., McInnes,L., Doyle,M., Hecksher-Sorensen,J. and Hastie,N.D. (1998) YAC transgenic analysis reveals Wilms' tumour 1 gene activity in the proliferating coelomic epithelium, developing diaphragm and limb. *Mech. Dev.*, **79**, 169–184.